

Reliability of 3 methods for assessing shoulder strength

Kimberley Hayes, BApSc(Physiotherapy),^{a,b} Judie R. Walton, PhD,^a Zoltan L. Szomor, MD,^a and George A. C. Murrell, MBBS, DPhil,^a Sydney, Australia

The reliability of tests for isometric strength of the shoulder joint in symptomatic subjects has yet to be established. For this purpose, interrater and intrarater agreement trials were undertaken to ascertain the reliability of manual muscle tests, a handheld dynamometer, and a spring-scale dynamometer for 5 different shoulder movements in symptomatic subjects. Intraclass correlation coefficients were calculated from a random-effects model. All movements tested with the handheld dynamometer demonstrated excellent reliability for the interrater trial ($\rho = 0.79-0.92$). Excellent reliability was also demonstrated for elevation, external rotation, and internal rotation for the intrarater trial ($\rho = 0.79-0.96$). For the interrater trial, measurement of the lift-off maneuver with the handheld dynamometer was significantly more reliable than with manual muscle tests ($P = .002$). In summary, the handheld dynamometer was the most reliable and discriminatory means for assessing strength of the rotator cuff in symptomatic subjects. (J Shoulder Elbow Surg 2002;11:33-9.)

INTRODUCTION

Disorders of the rotator cuff are the most common cause of shoulder pain and disability.^{2,34} Despite the prevalence of disorders of the rotator cuff and the many methods of managing them, little has been done to establish the reliability of an isometric strength assessment in patients with rotator cuff dysfunction.

Reliability is defined as the extent to which a measurement is repeatable²⁰ and can be estimated from measurements made by a single rater or multiple raters on the same material (agreement). Two different types of agreement can be distinguished according to whether one rater makes two or more measurements of

the same material (intrarater agreement) or each of several raters independently measures the same material (interrater agreement).

There have been many studies on the reliability of strength assessment tests* and the variables associated with measurement error.† However, the majority have been conducted on either subjects with normal shoulders^{3,5,6,8,11,17,32} or neurologically impaired patients.^{1,3,4,9,16,18,29,33} A limited number of studies have included patients with various orthopaedic conditions^{9,21,33}; however, in these cases the exact pathology has been poorly specified. To our knowledge, there are no published studies that have investigated the reliability of isometric strength assessment for the shoulder in patients with disorders of the shoulder.

The aim of this study was to determine the interrater and intrarater reliability of 3 tests for assessing isometric shoulder muscle strength in patients with a range of shoulder dysfunctions.

MATERIALS AND METHODS

Two separate groups of subjects gave informed consent for trials that investigated the interrater and intrarater reliability of 3 muscle strength assessment tests.

The interrater reliability trial consisted of 8 volunteers, 3 men and 5 women, ranging in age from 57 to 72 years (mean, 66 years). All subjects had a current shoulder complaint. Six patients had undergone rotator cuff repair surgery within the past 24 months, 1 patient was 17 months post scapulothoracic fusion, and 1 patient had adhesive capsulitis.

The intrarater reliability trial consisted of 9 volunteers, 5 men and 4 women, ranging in age from 29 to 74 years (mean, 64 years). One of these subjects had 2 symptomatic shoulders. Of the 10 shoulders, 8 were symptomatic, being within 36 months of rotator cuff repair surgery. The remaining 2 shoulders were asymptomatic at the time of testing.

Four raters, identified here as A to D, were used for the interrater reliability trial. Rater A was an orthopaedic surgeon, rater B was a sports physician trainee, and raters C and D were qualified physiotherapists. All raters tested all 8 subjects with the 3 muscle strength assessment tests. Each rater was blinded to the results of the other raters' assessments.

Only rater A was used for the intrarater reliability trial. This rater tested all 9 subjects (including the one subject being tested bilaterally) on 3 separate occasions within a

From the Department of Orthopaedic Surgery,^a St George Hospital Campus, University of New South Wales, Sydney, and School of Physiotherapy,^b Faculty of Health Sciences, University of Sydney, Australia.

Reprint requests: A/Prof George A. C. Murrell, Department of Orthopaedic Surgery, St George Hospital, Kogarah 2217 NSW Australia.

Copyright © 2002 by Journal of Shoulder and Elbow Surgery Board of Trustees.

1058-2746/2002/\$35.00 + 0 32/1/119852

doi:10.1067/mse.2002.119852

*References 1, 3-6, 8, 9, 11, 16-18, 21, 27, 29, 32, 33.

†References 10, 12, 13, 22, 23, 25, 26, 28, 30.



Figure 1 Handheld dynamometer used to test muscle strength of the shoulder.

48-hour period, using the same 3 muscle strength assessment tests.

The 3 tests used for measuring isometric muscle strength for both the interrater and intrarater trials comprised manual muscle tests, testing with a handheld dynamometer, and testing with a spring-scale dynamometer. Manual muscle tests were used for the movements of elevation, external rotation, internal rotation, and hand behind back lift-off. We graded muscle strength as follows: grade 1, trace or absent muscle contraction as determined by observation and palpation of the tendons and muscle bulk; grade 2, subject was able to move the affected extremity through its range of motion with gravity eliminated; grade 3, subject was able to move the affected extremity through its range of motion against the resistance of gravity; grade 4, subject had antigravity power and could hold a mid-range position against a low load resistance applied by the examiner (submaximal resistance for that individual); grade 4.5, subject had antigravity power and could hold a mid-range position against a greater resistance applied by the examiner (greater resistance than for grade 4 and lesser resistance than for grade 5); and grade 5, subject had antigravity power and could hold a mid-range position against a resistance that the examiner considered to be maximal for that individual.

A handheld dynamometer (PowerTrack MMT; JTech Medical Industries, Alpine, Utah, and Muscletester; Hoggan Health Industries, South Draper, Utah) was used for the movements of elevation, external rotation, internal rotation, and hand behind back lift-off. This device had a force-measuring capacity of 4.4 to 445 N, in 4.4-N increments (Figure 1).

A spring-scale dynamometer (Manley 2012 spring scale; Manley Tool and Machine, Independence, Mo) (Figure 2) was used for the movements of elevation, external rotation, internal rotation, and adduction. The lift-off maneuver was not measured with this device because it was not possible to perform this test with the spring-scale dynamometer. This device had a force-measuring capacity of 0 to 25 kg, in 0.25-kg increments. Velcro cuff loops were secured to either end of the spring scale to allow both subject and rater attachment. This device was tested as a



Figure 2 Spring-scale dynamometer used to test muscle strength of the shoulder.

possible cost-effective alternative to the handheld dynamometer.

For the interrater trial, all raters were briefed on the study protocol and a 5-hour training session was conducted to ensure familiarization and standardization of the 3 muscle strength tests. The intrarater trial was conducted several weeks after the interrater trial and involved a different group of subjects. Both trials followed the procedures outlined below.

For each test, raters were asked to ensure that resistance was applied perpendicular to the limb segment at the appropriate joint angle and that adequate manual stabilization was provided to prevent compensatory muscular action. A "make" test procedure^{6,7} was used for all 3 tests to determine the isometric strength for each movement. The subject was asked to build to a maximum contraction over a 1- to 2-second period and to hold the maximum effort against applied resistance for a further 4 to 5 seconds. The recorded measure reflected the maximum isometric value

achieved by the individual. Each rater performed all 3 tests on a given subject before starting the assessment of the next subject. This procedure was repeated until all 8 subjects had been measured by all 4 raters. With the exception of manual muscle tests, which were always performed first, the order of the other 2 tests was random. This was done to prevent inadvertent biasing of the manual muscle grade with objective data such as those provided by the handheld dynamometer or spring-scale dynamometer. The following assessment protocol for the 3 muscle strength tests was used.

(1) Manual muscle tests

(a) Elevation. The subject was seated upright on the edge of the treatment table with the feet supported on a footstool. The rater stood in front of the subject and placed the affected extremity into a position of 90° of elevation 30° in front of the coronal plane. The palm of the hand faced downward for this test. The subject was asked to hold this position as the rater applied a downward resistance through the distal end of the dorsal forearm (Figure 3).

(b) External rotation. The subject was positioned as for position 1a, and the rater stood to the subject's affected side. The subject's affected extremity was placed by his or her side with 90° elbow flexion and neutral forearm pronation/supination. He or she was asked to hold this position as the rater applied a medially directed resistance through the dorsal aspect of the distal forearm. The rater provided stabilization to the distal end of the humerus with the non-testing hand to prevent unwanted humeral movement.

(c) Internal rotation. The subject, the rater, and the affected extremity were positioned as for position 1b. The subject was asked to hold this position as the rater applied a laterally directed resistance through the volar aspect of the distal forearm. The rater provided stabilization to the distal end of the humerus with the non-testing hand to prevent unwanted humeral movement.

(d) Hand behind back lift-off maneuver. This maneuver was first described by Gerber and Krushell¹⁵ as a method with which to test subscapularis strength. The subject was positioned as for position 1a, and the rater stood behind him or her. The subject's affected extremity was placed midline behind the back to a reach that was governed by the individual's upper limb flexibility. The dorsal forearm was clear of skin contact. The subject was asked to hold this test position as the rater provided an anteriorly directed force through the volar aspect of the distal forearm.

(2) Handheld dynamometer

(a) Elevation. The subject, the rater, and the affected extremity were positioned as for position 1a. The dynamometer was centered on the dorsal aspect of the distal forearm and was kept parallel to the ground throughout the testing procedure. The subject was asked to hold this position as the rater applied a downward force through the dynamometer.

(b) External rotation. The subject was in the supine position, and the affected extremity was supported on the



Figure 3 Position used to assess elevation strength by manual muscle tests.

treatment table in 90° abduction, 90° elbow flexion, and a mid-forearm position. The rater stood to the affected side of the subject and stabilized the medial aspect of the distal humerus with the non-testing hand. The dynamometer was centered on the dorsal aspect of the distal forearm and was kept perpendicular to the ground throughout the testing procedure. The subject was asked to hold this position as the rater applied a caudally directed force through the dynamometer.

(c) Internal rotation. The subject, the rater, and the affected extremity were positioned as for position 2b. The rater stabilized the lateral aspect of the distal humerus with the non-testing hand. The dynamometer was centered on the volar aspect of the distal forearm and was kept perpendicular to the ground throughout the testing procedure. The subject was asked to hold this position as the rater applied a rostrally directed force through the dynamometer.

(d) Hand behind back lift-off maneuver. The subject, the rater, and the affected extremity were positioned as for position 1d. The dynamometer was centered on the volar aspect of the distal forearm and was kept perpendicular to the long axis of the forearm throughout the testing procedure. The subject was asked to hold the test position as the rater applied an anteriorly directed force through the dynamometer.

(3) Spring-scale dynamometer

(a) Elevation. The subject, the rater, and the affected extremity were positioned as for position 1a. The superior cuff of the spring-scale dynamometer was looped around the distal aspect of the subject's elevated forearm, and the subject was asked to hold this test position as the rater applied a downward force through the inferior cuff attachment. The body of the spring-scale dynamometer was kept perpendicular to the ground throughout the testing procedure.

(b) External rotation. The subject and the affected extremity were positioned as for position 1b. The rater stood in front of the subject. The superior cuff of the spring-scale dynamometer was looped around the distal aspect of the subject's affected forearm, and the subject was asked to hold this position as the rater applied a medially direct-

Table I Interrater reliability of 3 methods for assessing shoulder strength

	Intraclass correlation coefficient (ρ)	95% CI
Manual muscle test		
Elevation	0.72	0.38-0.93
External rotation	0.55	0.17-0.88
Internal rotation	0.61	0.26-0.89
Lift-off	0.38	0.02-0.81
Dynamometry		
Elevation	0.92	0.75-0.99
External rotation	0.82	0.55-0.96
Internal rotation	0.85	0.62-0.97
Lift-off	0.79	0.50-0.95
Spring-scale dynamometer		
Elevation	0.96	0.84-1.00
External rotation	0.75	0.40-0.95
Internal rotation	0.88	0.68-0.98
Adduction	0.90	0.72-0.98

Each test was performed on 8 subjects by 4 raters. Intraclass correlation coefficients (ρ) were calculated with use of a 2-way random-effects model.

ed force through the inferior cuff attachment. The body of the spring-scale dynamometer was kept parallel to the ground throughout the testing procedure.

(c) Internal rotation. The subject, the rater, and the affected extremity were positioned as for position 1b. The superior cuff of the spring-scale dynamometer was looped around the distal aspect of the subject's affected forearm, and the subject was asked to hold this position as the rater applied a laterally directed force through the inferior cuff attachment. The body of the spring-scale dynamometer was kept parallel to the ground throughout the testing procedure.

(d) Adduction. The subject stood, and the rater faced the affected side of the subject. The subject's affected extremity was abducted 30° from the neutral position. The superior cuff of the spring-scale dynamometer was looped around the distal aspect of the subject's affected extremity, and the subject was asked to hold this position as the rater applied a laterally directed force through the inferior cuff attachment.

For the interrater and intrarater reliability trials, 2-way random-effects intraclass correlation coefficients (ρ), together with confidence intervals, were calculated with SPSS (SPSS, Inc, Chicago, Ill) statistical software.³⁵ In accordance with the suggestions of Fleiss,¹⁴ a ρ value of less than 0.4 was deemed as representing poor reliability; 0.4 to 0.75 as fair to good reliability; and greater than 0.75 as excellent reliability.

In addition to determining the 2-way random-effects intraclass correlation coefficient, we calculated the Pearson correlation coefficient (r) for movements assessed with the handheld dynamometer and the spring-scale dynamometer. This correlation coefficient (r) was used as a measure of association for the 2 objective strength tests.

Table II Intrarater reliability of 3 methods for assessing shoulder strength

	Intraclass correlation coefficient (ρ)	95% CI
Manual muscle test		
Elevation	0.79	0.51-0.94
External rotation	0.86	0.66-0.96
Internal rotation	1.00	1.00-1.00
Lift-off	0.29	-0.08-0.71
Dynamometry		
Elevation	0.96	0.88-0.99
External rotation	0.92	0.78-0.98
Internal rotation	0.85	0.64-0.96
Lift-off	0.70	0.37-0.91
Spring-scale dynamometer		
Elevation	0.96	0.89-0.99
External rotation	0.84	0.61-0.95
Internal rotation	0.72	0.40-0.91
Adduction	0.89	0.71-0.97

Each test was performed on 9 subjects on 3 occasions by a single trained rater. Intraclass correlation coefficients (ρ) were calculated with use of a random-effects model.

RESULTS

The interrater intraclass correlation coefficients (ρ) for each movement tested with each of the strength tests are displayed in Table I. For all tests performed with the handheld dynamometer (elevation, external rotation, internal rotation, and lift-off) and for most tests performed with the spring-scale dynamometer (elevation, internal rotation, and adduction), agreement between the 4 raters was excellent ($\rho = 0.79-0.92$ and $\rho = 0.75-0.96$, respectively). Agreement for external rotation, as assessed by the spring-scale dynamometer, was fair to good ($\rho = 0.75$).

Correlation between the 2 types of dynamometer measurements was high, with Pearson correlation coefficients (r) of 0.88 for elevation, 0.77 for external rotation, and 0.82 for internal rotation ($P < .0001$).

For the 4 movements assessed by manual muscle tests (elevation, external rotation, internal rotation, and lift-off), agreement between the 4 raters ranged from poor to fair to good ($\rho = 0.38-0.72$). The lift-off maneuver, as assessed by manual muscle tests, was the least reproducible movement of the trial ($\rho = 0.38$). The same movement assessed by the handheld dynamometer was significantly more reliable than the manually tested lift-off maneuver ($\rho = 0.79$, $P < .002$).

Figure 4, A and B graphically represent the distribution of raw data for the lift-off maneuver as assessed by manual muscle tests and the handheld dynamometer, respectively. Interestingly, subjects 1 and 2 were assigned strength grades for the lift-off maneuver that ranged from grade 3 (no force-generating capacity against resistance) to grade 5 (maximal force-generating capacity) by raters using manual muscle tests (Fig-

ure 4, A). The variability was reduced when the handheld dynamometer was used (Figure 4, B).

The intrarater intraclass correlation coefficients (ρ) for each movement tested with each of the strength tests are displayed in Table II. For most tests performed with the handheld dynamometer (elevation, external rotation, and internal rotation) and for most tests performed with the spring-scale dynamometer (elevation, external rotation, and adduction), agreement between the 3 measurement occasions was excellent ($\rho = 0.85$ - 0.96 and $\rho = 0.84$ - 0.96 , respectively). Agreement for lift-off, as assessed by the handheld dynamometer, and internal rotation, as assessed by the spring-scale dynamometer, was fair to good ($\rho = 0.70$ and $\rho = 0.72$, respectively).

In the interrater trial, correlation between the 2 types of dynamometer measurements was high, with Pearson correlation coefficients (r) of 0.99 for elevation, 0.85 for external rotation, and 0.86 for internal rotation ($P < .0001$).

For the 4 movements assessed by manual muscle tests, agreement between the 3 measurement occasions was excellent for elevation ($\rho = 0.79$), external rotation ($\rho = 0.86$), and internal rotation ($\rho = 1.00$), but poor for the lift-off maneuver ($\rho = 0.29$).

Data for internal rotation measured by the spring-scale dynamometer was converted from kilograms to Newtons of force, to allow for direct comparison with the handheld dynamometer. For this movement, all subjects were assigned grade 5 muscle strength, by means of manual muscle tests. In contrast, the same subjects effected strength outcomes that ranged from 50 to 175 N with the handheld dynamometer and 49 to 196 N with the spring-scale dynamometer.

DISCUSSION

The results of this study show that the handheld dynamometer and the spring-scale dynamometer are reliable methods for assessing isometric shoulder muscle strength in subjects with shoulder joint disease. The reliability of manual muscle tests was less consistent than that of the handheld dynamometer and the spring-scale dynamometer, as demonstrated by correlation coefficients that were as low as 0.29.

Manual muscle tests are the most widely used method of physical examination for clinical evaluation of muscle strength.²⁵ However, manual muscle tests are also recognized as being a subjective measurement of strength, dependent on the experience, strength, and judgment of the examiner.^{1,3,21,25,27,29,33} In assigning an ordinal grade to the test muscle group, the examiner relies on an internal basis for comparing test results with what he or she considers to be normal for age, sex, and weight of the individual.³³ The results of our study support the argument of individual examiner bias in the use of manual muscle tests. This was well demonstrated in the interrater trial (Figure 4, A) for the lift-off maneuver, in which raters judged subjects 1 and 2 as having lift-off strength that ranged from no force-gener-

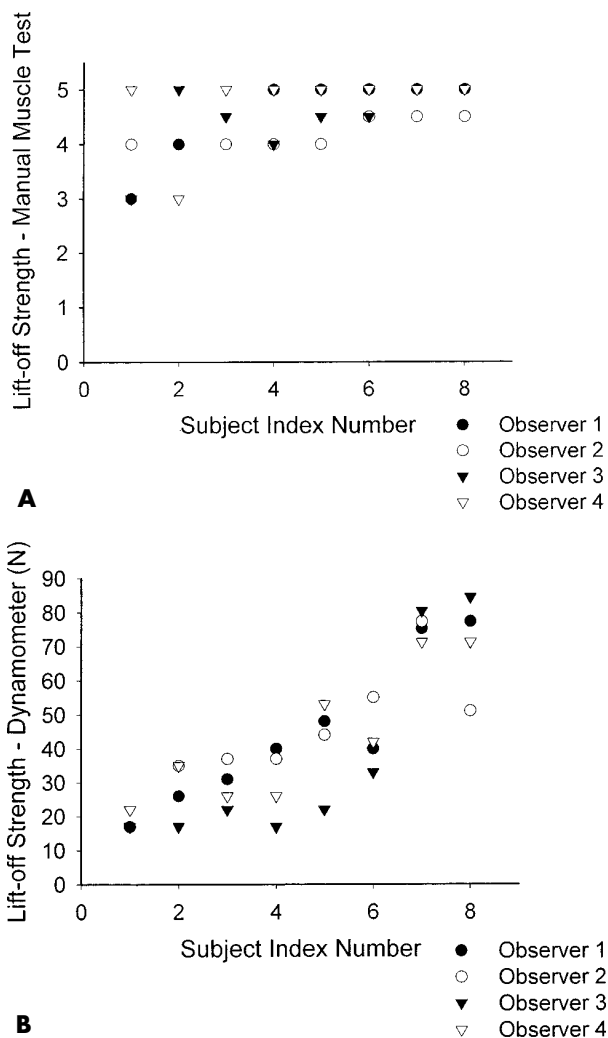


Figure 4 Variations in lift-off strength measurements made by 4 raters in 8 subjects for manual muscle tests (A) and handheld dynamometer (B).

ating capacity against manual resistance (grade 3) to maximal force-generating capacity (grade 5). Further, manual muscle tests generally had better intrarater reliability than interrater reliability. These results suggest that an examiner is reliable in reproducing a muscle grade in accordance with his or her own grading system, but less able to reproduce a second examiner's outcome under similar measurement conditions.

A further criticism of manual muscle tests has been their inability to detect small differences in muscle strength, particularly in the good to normal strength categories.^{3,21,29,34} This study confirmed the lack of discrimination between various strength outcomes when assessed by manual muscle tests. The ability of manual muscle tests to discriminate between internal rotation power was particularly poor. For the intrarater trial,

internal rotation strength values varied from 50 to 175 N with the handheld dynamometer and from 49 to 196 N with the spring-scale dynamometer. However, the same subjects were all deemed to have grade 5 strength when assessed by manual muscle tests. Thus the information obtained from manual muscle tests appears to be relatively imprecise and of less value in identifying changes in strength status.

In 1991 Gerber and Krushell¹⁵ described a lift-off test with which to evaluate the integrity of the subscapularis musculotendinous unit. In that study, patients with a subscapularis rupture were unable to effect the maneuver. More recently, electromyographic analysis of selected shoulder girdle muscles was conducted for 4 versions of the lift-off test to establish a precise starting position for the upper limb that would best isolate the internal rotation function of the subscapularis muscle from other synergists to this movement.³¹ For most versions of this test, the subscapularis muscle was shown to be highly active. However, a position of maximum internal rotation whereby the dorsum of the hand was placed in the midline of the posterior thorax at the level of the inferior aspect of the scapulae was shown to be the position best able to isolate the internal rotation function of subscapularis.³¹ For patients with supraspinatus tears, it may be particularly difficult to achieve this position, as this patient group has reduced range of active internal rotation.²⁴ To our knowledge, the reliability of this test has not been determined.

Raters were poor at manually assessing strength in this position in these patients. When the same movement was measured with the handheld dynamometer, reproducibility was considerably improved ($P = .002$ in the interrater trial). By using a measurement test that provides a direct recording of force output, clinicians may be able to assess the integrity of the subscapularis complex with greater accuracy than that afforded by a subjective grading system.

The results of this study are in keeping with those of previous trials that have shown the handheld dynamometer to be a reliable measurement tool, from both an interrater and intrarater perspective, for the measurement of shoulder muscle strength in patients with normal shoulders and neurologically impaired patients.^{4,8,9,32} Intrarater reliability coefficients for various shoulder muscle groups have been stated to be as high as 0.95 to 0.99,⁴ 0.98,³² and 0.95 to 0.97,⁸ to a somewhat lower value of 0.69.³³ Interrater reliability of the handheld dynamometer at the shoulder joint has not been examined as extensively. A value of 0.88 has been reported for shoulder external rotation,⁹ but to our knowledge, interrater reliability for other shoulder movements has not been reported.

Although its reliability has not been investigated as extensively as that of other strength assessment tests, there is evidence to suggest that reproducible measurements can be taken with a spring-scale dynamome-

ter.^{10,19} The results of this study demonstrated excellent reliability for most movements assessed with the spring-scale dynamometer. In addition, high correlation was demonstrated for measurements taken with the spring-scale dynamometer and the handheld dynamometer. To its advantage, the spring-scale dynamometer is portable, is easy to operate, and is a cost-effective measurement tool (assembly cost, approximately \$200). However, assessment of the lift-off maneuver was biomechanically impossible with the spring-scale dynamometer, because the instrument measured distraction force rather than compression force, which limits its usefulness in the assessment of subscapularis muscle strength.

In summary, reliable isometric strength measurement of the shoulder musculature in symptomatic subjects was shown with the handheld dynamometer and the spring-scale dynamometer. The reliability of manual muscle tests was found to be less consistent than that of the handheld dynamometer and the spring-scale dynamometer. Lift-off strength was unreliable when assessed manually. Thus, for a strength examination that requires information specific to all components of the rotator cuff, the handheld dynamometer is the strength assessment test of choice.

We wish to thank the patients and clinicians [Larry Bryant, MBBS, BSc, Kim Wade, BApSc(Physiotherapy), and Sophia Short, BApSc(Physiotherapy)] who assisted in this study. We also thank Robert Crouch (Precision Fitness Industries, Hindmarsh, Adelaide, South Australia) and Karen Ginn (Senior Lecturer, Department of Biological Sciences, Health Sciences Faculty, University of Sydney) and acknowledge the support of the St George Private Hospital/Health Care of Australia and the Sutherland Hospital, South Eastern Sydney Area Health Service.

REFERENCES

1. Aitkens S, Lord J, Bernauer E, Fowler WM, Lieberman JS, Berck P. Relationship of manual muscle testing to objective strength measurements. *Muscle Nerve* 1989;17:3-7.
2. Bartolozzi A, Andreychik D, Ahmad S. Determinants of outcome in the treatment of rotator cuff disease. *Clin Orthop* 1994;308:90-7.
3. Beasley WC. Influence of method on estimates of normal knee extensor force among normal and postpolio children. *Phys Ther Rev* 1956;36:21-41.
4. Bohannon RW. Test-retest reliability of hand-held dynamometry during a single session of strength assessment. *Phys Ther* 1986;66:206-9.
5. Bohannon RW. Hand-held dynamometry: stability of muscle strength over multiple measurements. *Clin Biomech* 1987;2:74-7.
6. Bohannon RW. Make tests and break tests of elbow flexor muscle strength. *Phys Ther* 1988;68:193-4.
7. Bohannon RW. Testing isometric limb muscle strength with dynamometers. *Crit Rev Phys Rehabil Med* 1990;2:75-86.
8. Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20-79 years. *Arch Phys Med Rehabil* 1997;78:26-32.
9. Bohannon RW, Andrews AW. Inter-rater reliability of hand-held dynamometry. *Phys Ther* 1987;67:931-3.

10. Bohannon RW, Andrews AW. Accuracy of spring and strain gauge hand-held dynamometers. *J Sports Phys Ther* 1989;10:323-5.
11. Bohannon RW, Saunders N. Hand-held dynamometry: a single trial may be adequate for measuring muscle strength in healthy individuals. *Physiother Can* 1990;42:6-9.
12. Clarke HH, Elkins EC, Wakim KG. Relationship between body position and the application of muscle power to movements of the joints. *Arch Phys Med Rehabil* 1950;31:81-9.
13. Figoni SF, Morris AF. Effects of knowledge of results on reciprocal, isokinetic strength and fatigue. *J Orthopaed Sports Phys Ther* 1984;6:190-7.
14. Fleiss JL. Reliability of measurement. In: *The design and analysis of clinical experiments*. New York: John Wiley & Sons; 1986. p.1-32.
15. Gerber C, Krushell RJ. Isolated rupture of the tendon of the subscapularis muscle. *J Bone Joint Surg Br* 1991;73:389-94.
16. Iddings DM, Smith LK, Spencer WA. Muscle testing: part 2. Reliability in clinical use. *Phys Ther Rev* 1961;41:249-56.
17. Jones RE. Reliability of muscle strength testing under varying motivational conditions. *J Am Phys Ther Assoc* 1962;42:240-3.
18. Lilienfeld AM, Jacobs M, Willis M. A study of the reproducibility of muscle testing and certain other aspects of muscle scoring. *Phys Ther Rev* 1954;34:279-89.
19. Lovett RW, Martin EG. Certain aspects of infantile paralysis (with a description of a method of muscle testing). *JAMA* 1916;66:729-33.
20. Lwanga SK, Tye C-Y, editors. *Teaching health statistics. Twenty lesson and seminar outlines*. Geneva: World Health Organization; 1986. p. 15.
21. Marino M, Nicholas JA, Gleim GW, Rosenthal P, Nicholas SJ. The efficacy of manual assessment of muscle strength using a new device. *Am J Sports Med* 1982;10:360-4.
22. McGarvey SR, Morrey BF, Askew IJ, An K. Reliability of isometric strength testing—temporal factors and strength variation. *Clin Orthop* 1984;185:301-5.
23. Mendler HM. Effect of stabilization on maximum isometric knee extensor force. *Phys Ther* 1967;47:375-9.
24. Murrell GA, Walton J. Diagnosis of rotator cuff tears. *Lancet* 2001;357:769-70.
25. Nicholas JA, Sapega A, Kraus H, Webb JN. Factors influencing manual muscle tests in physical therapy. *J Bone Joint Surg Am* 1978;60:186-90.
26. Peacock B, Westers T, Walsh S, Nicholson K. Feedback and maximum voluntary contraction. *Ergonomics* 1981;24:223-8.
27. Saraniti AJ, Gleim GW, Melvin M, Nicholas JA. The relationship between subjective and objective measurements of strength. *J Orthopaed Sports Phys Ther* 1980;2:15-9.
28. Schenck JM, Forward EM. Quantitative strength changes with test repetitions. *J Am Phys Ther Assoc* 1965;45:562-9.
29. Schwartz S, Cohen ME, Herbison GJ, Shah A. Relationship between two measures of upper extremity strength: manual muscle test compared to hand-held myometry. *Arch Phys Med Rehabil* 1992;73:1063-8.
30. Smidt GL, Rogers MW. Factors contributing to the regulation and clinical assessment of muscular strength. *Phys Ther* 1982;62:1283-90.
31. Stefko JM, Jobe FW, VanderWilde RS, Carden E, Pink M. Electromyographic and nerve block analysis of the subscapularis lift-off test. *J Shoulder Elbow Surg* 1997;6:347-55.
32. Sullivan J, Chesley A, Hebert G, McFaul S, Scullion D. The validity and reliability of hand-held dynamometry in assessing isometric external rotator performance. *J Orthopaed Sports Phys Ther* 1988;10:213-7.
33. Wadsworth CT, Krishnan R, Sear M, Harrold J, Nielsen DH. Intrarater reliability of manual muscle testing and hand-held dynamometric muscle testing. *Phys Ther* 1987;67:1342-7.
34. Wirth MA, Basamania C, Rockwood CA. Non operative management of full-thickness tears of the rotator cuff. *Orthop Clin North Am* 1997;28:59-67.
35. Yaffee RA. Enhancement of reliability analysis: application of intraclass correlations with SPSS/Windows v.8. 1998. Available at: <http://www.nyu.edu/acf/socsci/Docs/intracls.html>. Accessed October 3, 2001.